2023-24 post-collection outputs

2023-24 Student numbers technical document and rebuild instructions

Contents

Summary

1. This document describes the algorithms applied to the 2023-24 Student (23056) record and Individualised Learner Record (ILR) data to determine a provider's higher education and further education student numbers. These numbers will be used by the Office for Students (OfS) for regulatory purposes such as calculating OfS registration fees, assessing applications for degree awarding powers and university title and determining mandatory participation in the Teaching Excellence Framework (TEF).

2. The student numbers for regulatory purposes will be released to providers as part of the 2023-24 post-collection outputs. The post-collection outputs will be based on the data originally submitted by the provider and signed off as accurate by the accountable officer or a provider's amended data provided it was submitted and approved by 7 May 2025. We use amended data submitted after 7 May 2025 for these purposes only when this use has been formally approved by the OfS.

3. The student numbers post collection outputs (PCO) are formed of two files. These are:

    a. Student numbers workbook 'SN23_PCO_XXXXXXXX.xlsx'

    b. Student numbers individualised file 'SN23_PCO_XXXXXXXX_IND.csv'

    where XXXXXXXX is the UK Provider Reference Number (UKPRN) for the provider.

4. Some providers submit data to both the OfS's designated data body (DDB) and the Education and Skills Funding Agency (ESFA). Student engagements meeting either of the criteria listed below should be reported to both:

    a. Students on higher and degree apprenticeships for whom you receive direct funding from the ESFA (PROGTYPE 20, 21, 22, 23 or 25).

    b. Students on higher education engagements (Level 4 or above) who are financed by Advanced Learner Loans.

    We would expect all other student engagements to be returned to either the DDB or the ESFA, but not both. For those providers that return both the DDB and ILR data, the student engagements listed above will be excluded from their ILR data to avoid double counting.

5. This document is aimed at readers with in-depth knowledge of the data. Readers are advised to have a copy of the Student record coding manual and ILR specification for the relevant year, and the Office of Qualifications and Examinations Regulation (Ofqual) Register Data Extract Specification to hand when using this document. Throughout the document, fields taken or derived from the Student (23056) or ILR records are shown in capitals.

6. We will document any changes or corrections to the algorithms in this document on the OfS website (www.officeforstudents.org.uk/data-and-analysis/student-number-data/).

Changes

Changes from 2022-23 the student numbers method

7. This section describes any changes to the algorithms from the 2022-23 student numbers method. Any changes after initial publication will be listed here.

Changes from the 2022-23 DDB student numbers method

PRINONUK

8. PRINONUK has been renamed to ENGPRINONUK and moved from the Study Location entity to the Engagement entity.

SNLEVELNUM

9. Removal of values 'D0005' and 'E0005' as these are no longer valid values for 2023-24.

Value	Description	Definition
8	Records at Level 8	Z_LEVEL = D0003, D0004, E0000, E0001, E0002, E0003, E0004, L0000
7	Records at Level 7	Z_LEVEL = L0001, L0002, L0003, ~~D0005~~, M0002, M0003, M0004, M0006, M0007, M0008, M0009, M0010, M0011, M0012, M0013, M0015, M0016, M0017, M0018, M0020, M0021, M0022, M0023, M0024, ~~E0005~~
6	Records at Level 6	Z_LEVEL = H0003, H0004, H0005, H0006, H0007, H0008, H0009, H0010, H0012, H0013, H0014, H0015, H0016, H0018, H0019, H0020, I0001
5	Records at Level 5	Z_LEVEL = I0002, I0004, I0005, I0006, I0007, I0008 I0009, I0010, I0012, I0013, J0000, J0001, J0002, J0003, J0004, J0005, J0006, J0007, J0010, J0011, J0012
4	Records at Level 4	Otherwise

Changes from the 2022-23 ILR student numbers method

10. There have been no changes to the algorithms.

Note on terminology

11. The algorithms below apply to two different data collections, each of which has a different structure. In the Student (23056) return, the fields that we use are associated with a student engagement. In the ILR return, the fields that we use are associated with a course.

12. Below, we use the generic term 'record' to refer to either a student engagement or a learner-course combination, depending on the data source.

13. The year '20Y1-Y2' below refers to the academic year starting on 1 August 2023 and ending on 31 July 2024.

Student numbers workbook

14. The student numbers workbook is the Excel workbook SN23_PCO_XXXXXXXX.xlsx which contains the following worksheets:

Table 1: Worksheets in the student numbers workbook

Worksheet*	Title
Information	Information sheet
Student Numbers	Number of student FTEs by broad level and HE study level

* This worksheet reference corresponds to the spreadsheet tabs.

Student numbers individualised file

15. When working through this document it is necessary to use the individualised file, SN23_PCO_XXXXXXXX_IND.csv. The individualised file contains the values of all fields used to generate student numbers.

16. Fields taken from the individualised student return or derived as part of the student numbers outputs are shown in capitals using the names given in Tables 2, 3 and 4 .

17. Specific instructions for rebuilding the student numbers workbook from the individualised file are provided in this document.

20Y1-Y2 fields used to create the student number outputs

18. The fields used to generate the student numbers outputs are listed in Tables 2, 3 and 4.

Table 2: Student fields used in the student number calculations

Entity	Name	Description
Engagement	UKPRN	UK Provider Reference Number
Engagement	SID	Student identifier
Engagement	NUMHUS	Engagement number
Engagement	ENGSTARTDATE	Engagement start date
Engagement	ENGENDDATE	Engagement end date
Engagement	INCOMINGEXCHANGE	Identifies if the student is an incoming exchange or visiting student
Engagement	ENGPRINONUK	Identifies if the student is spending the majority of their Engagement outside the UK
Course	PREREQUISITE	Identifies PGR courses for funding purposes.
Engagement derived field	Z_ACT_CYC	Marker for student active (includes writing-up) within the cycle
Engagement derived field	Z_LEVEL	Student's course level for the engagement
Engagement derived field	Z_STULOAD_CYC	Total student load for the engagement across the cycle

Table 3: ILR, Learning Aim Reference Service and Ofqual fields used in the student number calculations

Source	Field	Description
ILR	COMPSTATUS	Degree of completion of the learning activities leading to the learning aim
ILR	LEARNACTENDDATE	Learning actual end date
ILR	LEARNPLANENDDATE	Learning planned end date
ILR	LEARNSTARTDATE	Learning start date
ILR	STULOAD	Learner FTE
ILR	PLANLEARNHOURS	Planned learning hours
ILR	PRIORLEARNFUNDADJ	The proportion of this learning aim still to be delivered
ILR	PROGTYPE	The type of programme which the learner is undertaking
ILR	UKPRN	UK Provider Reference Number
ILR	WITHDRAWREASON	The reason why the learner has withdrawn from the learning aim
LARS	LEARNAIMREF	Learning aim reference
ILR	AIMSEQNUMBER	Learning aim data set sequence
LARS	LEARNAIMREFTITLE	Learning aim title
LARS	LEARNAIMREFTYPE	Learning aim type
LARS	NVQ_LV	Notional NVQ level
LARS	NOTIONALNVQLEVELV2	The detailed learning aim level on the NVQ scale
Ofqual	TOTAL QUALIFICATION TIME	The total time a learner may typically be expected to expend in achieving a qualification
Ofqual	MAXIMUM GUIDED LEARNING HOURS	The maximum number of hours a learner can be expected to be taught or instructed by their tutor or teacher to complete the qualification
Ofqual	TOTAL CREDITS	The total credits required to obtain the qualification

Description of derived fields

19. The derived fields used to generate the student numbers outputs are listed in Table 4.

Table 4: Description of derived fields for recreating tables

Derived field name	Description	Paragraph
SNSOURCE	Identifies the source of the record	28
LEARNDELFAM_ADL1	Advanced Learner Loan indicator	30
SNDDBILRFLAG	Identifies providers that return both Student and ILR data	31
SNEXCL1 – SNEXCL256	Student number population exclusions	32- 40
SNEXCL	Reason(s) for exclusion from student numbers population	41
OFSQAIM[1]	Broad category of higher education	XX
SNLEVEL	Identifies record as higher or further education	43
SNLEVELNUM	Level of higher education record	44
SNTITLEHRS	Max learning hours in course title	45
SNTITLECREDITS	Max credits in course title	46
SNPRIORLEARNADJ	Adjustment for prior learning	47
SNQUALHOURS	Total hours course is expected to take	48
SNENDDATE	Last day studied in academic year	49
SNAYDAYSSTUDIED	Number of days studied in academic year	50
SNAVHOURSPERDAY	Average study hours per day over course	52
SNHOURSPERAYR	Total hours studied in academic year	54
SNHESTULOAD	Student load for higher education record	55
SNHEENGLEVSTULOAD	Student load per engagement (for DDB data) or level (for ILR data)	56
SNHEENGLEVSTULOADADJ	Capped student load per engagement (for DDB data) or level (for ILR data)	57
SNTOTHESTULOAD	Higher education student load per student	58
SNTOTHEFTE	Higher education FTE per student	59
SNFEPLANLEARNSTULOAD	Further education student load calculated from PLANLEARNHOURS	60
SNNUMFERECORDS	Number of non-excluded further education records per student	61
SNFESTULOAD	Student load for further education record	62

---

[1] OFSQAIM is derived using the algorithms documented in 'HESES24 Course table information', available on the OfS website at www.officeforstudents.org.uk/data-and-analysis/data-collection/heses/.

SNTOTFESTULOAD	Further education student load per student	63
SNTOTFEFTE	Further education FTE per student	64
SNREL_TEF	Undergraduate higher education provision relevant to TEF	65
SNHEENGLEVSTULOAD_TEF	Student load per engagement (for DDB data) or level (for ILR data) relevant to the Teaching Excellence Framework (TEF)	66
SNHEENGTLEVSTULOADADJ_TEF	Capped student load per engagement (for DDB data) or level (for ILR data) relevant to TEF	67
SNTOTHESTULOAD_TEF	Higher education student load relevant to TEF per student	68
SNTOTHEFTE_TEF	Higher education FTE relevant to TEF per student	69

Derived fields for calculating total higher education FTE

SNSOURCE

20. This field identifies the source of the record.

Value	Definition
DDB	Record taken from the Student record
ILR	Record taken from the ILR

21. Where an algorithm cannot be applied in the same way to each SNSOURCE, this will be indicated in the description of the algorithm.

LEARNDELFAM_ADL1

22. For ILR records only. This field contains the value of LEARNDELFAMCODE where LEARNDELFAMTYPE = ADL.

SNDDBILRFLAG

23. This field identifies providers that returned Student data to the OfS's DDB and ILR data in the academic year 20Y1-Y2.

Value	Definition
1	Provider returned both DDB and ILR data for the academic year 20Y1-Y2
0	Provider did not return both DDB and ILR data for the academic year 20Y1-Y2

SNEXCL1

24. This field identifies records that start after the academic year of interest.

Value	Description	Definition
1	Excluded	(SNSOURCE = DDB and ENGSTARTDATE > 31 July 20Y2 or (SNSOURCE = ILR and LEARNSTARTDATE > 31 July 20Y2)
0	Included	Otherwise

SNEXCL2

25. This field identifies records that finish before the academic year of interest.

Value	Description	Definition
1	Excluded	(SNSOURCE = DDB and ENGENDDATE < 01 August 20Y1 and ENGENDDATE is not BLANK) or (SNSOURCE = ILR and LEARNACTENDDATE < 01 August 20Y1 and LEARNACTENDDATE is not BLANK)
0	Included	Otherwise

SNEXCL4

26. This field identifies records that have been inactive for the reporting year.

Value	Description	Definition
1	Excluded	SNSOURCE = DDB and Z_ACT_CYC = 0
0	Included	Otherwise

SNEXCL8

27. This field identifies incoming exchange students and students who are primarily outside the UK for their engagement.

Value	Description	Definition
1	Excluded	SNSOURCE = DDB and (INCOMINGEXCHANGE ≠ BLANK or ENGPRINONUK = 01)
0	Included	Otherwise

SNEXCL16

28. This field identifies apprenticeship standard 'wrapper' programme aims.

Value	Description	Definition
1	Excluded	SNSOURCE = ILR and LEARNAIMREF = ZPROG001
0	Included	Otherwise

SNEXCL32

29. This field identifies higher apprentices in the ILR that have been returned by providers that also submit DDB data.

Value	Description	Definition
1	Excluded	SNSOURCE = ILR and SNDDBILRFLAG = 1 and PROGTYPE = 20, 21, 22, 23, 25
0	Included	Otherwise

SNEXCL64

30. This field identifies a record which has been closed to correct an incorrect LEARNPLANENDDATE.

Value	Description	Definition
1	Excluded	SNSOURCE = ILR and COMPSTATUS = 4[2] and WITHDRAWREASON = 40 and LEARNACTENDDATE = 01 August 20Y1
0	Included	Otherwise

SNEXCL128

31. This field identifies records which we believe relate to teaching provision rather than registered provision. This will be identified if two records returned by two providers for the same student are found to have the same LEARNAIMREF, the courses start within 31 days of each other and the UKPRN for one record matches the PARTNERUKPRN of the other record. The record with missing PARTNERUKPRN will then be flagged.

Value	Description	Definition
1	Excluded	SNSOURCE = ILR and the student appears to exist in another provider's ILR.
0	Included	Otherwise

---

[2] ILR records returned to the ESFA with COMPSTATUS=3 and WITHDRAWREASON=40 will have the value of COMPSTATUS changed to 4 by the ESFA. This is replicated in the ILR post collection output.

SNEXCL256

32. This field identifies students on higher education engagements financed by an Advanced Learner Loan that have been returned by providers that also submit Student (23056) data.

Value	Description	Definition
1	Excluded	SNSOURCE = ILR and SNDDBILRFLAG = 1 and LEARNDELFAM_ADL1 = 1 and NOTIONALNVQLEVELV2 = 4, 5, 6
0	Included	Otherwise

SNEXCL

33. This field indicates whether the record will be included in the calculation of student numbers.

34. This field contains the exclusion reason(s) for the engagement. It is calculated as (1 × SNEXCL1) + (2 × SNEXCL2) + …+ (256 × SNEXCL256).

SNLEVEL

35. This field identifies whether the record is of higher education or further education.

Value	Description	Definition
HE	Record of higher education	SNSOURCE = DDB or (SNSOURCE = ILR and LEARNAIMREF ≠ ZPROG001 and NOTIONALNVQLEVELV2 = 4, 5, 6, 7, 8, H)
FE	Record of further education	LEARNAIMREF ≠ ZPROG001 and not above
BLANK	Apprenticeship standard 'wrapper' programme aim	LEARNAIMREF=ZPROG001

SNLEVELNUM

36. This field indicates the level of study for higher education records, using the sector recognised standards as set out in condition of registration B5[3].

SNSOURCE = DDB

Value	Description	Definition
8	Records at Level 8	Z_LEVEL = D0003, D0004, E0000, E0001, E0002, E0003, E0004, L0000
7	Records at Level 7	Z_LEVEL = L0001, L0002, L0003, M0002, M0003, M0004, M0006, M0007, M0008, M0009, M0010, M0011, M0012, M0013, M0015, M0016, M0017, M0018, M0020, M0021, M0022, M0023, M0024

---

[3] See www.officeforstudents.org.uk/publications/sector-recognised-standards/.

6	Records at Level 6	Z_LEVEL = H0003, H0004, H0005, H0006, H0007, H0008, H0009, H0010, H0012, H0013, H0014, H0015, H0016, H0018, H0019, H0020, I0001
5	Records at Level 5	Z_LEVEL = I0002, I0004, I0005, I0006, I0007, I0008 I0009, I0010, I0012, I0013, J0000, J0001, J0002, J0003, J0004, J0005, J0006, J0007, J0010, J0011, J0012
4	Records at Level 4	Otherwise

SNSOURCE = ILR

Value	Description	Definition
8	Records at Level 8	SNLEVEL = HE and NOTIONALNVQLEVELV2 = 8
7	Records at Level 7	SNLEVEL = HE and NOTIONALNVQLEVELV2 = 7
6	Records at Level 6	SNLEVEL = HE and NOTIONALNVQLEVELV2 = 6
5	Records at Level 5	SNLEVEL = HE and NOTIONALNVQLEVELV2 = 5
4	Records at Level 4	SNLEVEL = HE and NOTIONALNVQLEVELV2 = 4, H
BLANK	Records not Level 4-8	Otherwise

SNTITLEHRS

37. For ILR records only. This field identifies the maximum number of hours per record as extracted from the course title on the Learning Aim Reference Service (LARS). For example, the value of SNTITLEHRS for the course listed as 'Non regulated SFA formula funded provision, Entry level, Medicine and Dentistry, 389 to 580 hrs, PW B' would be 580.

SNTITLECREDITS

38. For ILR records only. This field identifies the maximum number of credits per record as extracted from the course title on the LARS. For example, the value of SNTITLECREDITS for the course listed as 'Innovation code, Entry level, Medicine and Dentistry, 3-5 credits' would be 5.

SNPRIORLEARNADJ

39. For ILR records only. This field identifies the funding adjustment for prior learning expressed as a decimal value.

Value	Definition
PRIORLEARNFUNDADJ/100	PRIORLEARNFUNDADJ > 0
0	PRIORLEARNFUNDADJ = 0

SNQUALHOURS

40. For ILR records only. This field determines the total number of hours the course associated with the record is expected to take.

Value	Description	Definition
TOTAL QUALIFICATION TIME × SNPRIORLEARNADJ	Source of hours is TOTAL QUALIFICATION TIME	STULOAD = BLANK and TOTAL QUALIFICATION TIME ≠ BLANK, 0
MAXIMUM GUIDED LEARNING HOURS × SNPRIORLEARNADJ	Source of hours is MAXIMUM GUIDED LEARNING HOURS	STULOAD = BLANK and MAXIMUM GUIDED LEARNING HOURS ≠ BLANK, 0 and not above
(TOTAL CREDITS × 10) × SNPRIORLEARNADJ	Source of hours is TOTAL CREDITS	STULOAD = BLANK and TOTAL CREDITS ≠ BLANK ,0 and not above
(SNTITLEHRS) × SNPRIORLEARNADJ	Source of hours is SNTITLEHRS	STULOAD = BLANK and SNTITLEHRS ≠ BLANK and not above
(SNTITLECREDITS × 10) × SNPRIORLEARNADJ	Source of hours is SNTITLECREDITS	STULOAD = BLANK and SNTITLECREDITS ≠ BLANK and not above
BLANK	Unknown qualification hours	Otherwise

SNENDDATE

41. For ILR records only. This field determines the end date to be used when calculating the number of days studied (SNAYDAYSSTUDIED) in the academic year 20Y1-Y2. Used only where SNQUALHOURS will be used to calculate the FTE of the record.

Value	Description	Definition
LEARNACTENDDATE	Learner completed learning activities in 20Y1-Y2	LEARNACTENDDATE ≠ BLANK and LEARNACTENDDATE <= 31 July 20Y2
31 July 20Y2	Learner is still studying at the end of 20Y1-Y2	Otherwise

SNAYDAYSSTUDIED

42. For ILR records only. This field determines the number of days studied in the 20Y1-Y2 academic year.

43. Calculated as the numbers of days between SNENDDATE and either 1 August 20Y1 or LEARNSTARTDATE, whichever is later. Used only where SNQUALHOURS will be used to calculate the FTE of the record.

SNAVHOURSPERDAY

44. For ILR records only. This field calculates the average number of hours that would be studied each day for the duration of the course.

45. Calculated as SNQUALHOURS divided by the number of days between the LEARNSTARTDATE and LEARNPLANENDDATE. Used only where SNQUALHOURS will be used to calculate the FTE of the record.

SNHOURSPERAYR

46. For ILR records only. This field calculates the total number of hours studied in the academic year for each record. Used only where SNQUALHOURS will be used to calculate the FTE of the record. We cap SNHOURSPERAYR at the value of SNQUALHOURS for the course.

- SNHOURSPERAYR = min ((SNAVHOURSPERDAY × SNAYDAYSSTUDIED), SNQUALHOURS).

SNHESTULOAD

47. This field determines the STULOAD for each higher education record. When using DDB data each record relates to an engagement. When using ILR data, a record refers to a learner-course combination.

Value	Definition
Z_STULOAD_CYC	SNSOURCE=DDB and SNEXCL=0
STULOAD	SNSOURCE=ILR and SNEXCL = 0 and SNLEVEL = HE and STULOAD ≠ BLANK
(SNHOURSPERAYR ÷ 540) × 100	SNSOURCE=ILR and SNEXCL = 0 and SNLEVEL = HE and STULOAD = BLANK and SNQUALHOURS ≠ BLANK
25	SNSOURCE=ILR and SNEXCL = 0 and SNLEVEL = HE and SNQUALHOURS= BLANK
0	SNEXCL > 0 and SNLEVEL = HE
BLANK	Otherwise

SNHEENGLEVSTULOAD

48. This field determines the total STULOAD for each higher education engagement.[4]

Value	Definition
SNHESTULOAD	SNSOURCE=DDB
Sum of SNHESTULOAD across all records for each unique combination of student (LEARNREFNUMBER) and level (SNLEVELNUM)	SNSOURCE=ILR

SNHEENGLEVSTULOADADJ

49. This field calculates the capped student load for each higher education engagement[5] .

$$SNHEENGLEVSTULOADADJ = min(SNHEENGLEVSTULOAD, 150).$$

SNTOTHESTULOAD

50. This field is the total student load per student including any capping, calculated as the sum of SNHEENGLEVSTULOADADJ per student.

SNTOTHEFTE

51. This field calculates the total higher education FTE per student including any capping, calculated as SNTOTHESTULOAD÷100.

Additional fields required to calculate total further education FTE

SNFEPLANLEARNSTULOAD

52. For ILR records only. This field converts PLANLEARNHOURS to student load for calculation of further education FTE.

- SNFEPLANLEARNSTULOAD = max(((PLANLEARNHOURS ÷ 540) × 100) - SNTOTHESTULOAD, 0).

SNNUMFERECORDS

53. For ILR records only. This field identifies the number of records per student with SNLEVEL = FE and SNEXCL = 0.

---

[4] The ILR data does not include the concept of an 'engagement'. For the purpose of calculating SNHEENGLEVSTULOAD we treat a level of study in the ILR data as analogous to an engagement.

[5] The ILR data does not include the concept of an 'engagement'. For the purpose of calculating SNHEENGLEVSTULOADADJ we treat a level of study in the ILR data as analogous to an engagement.

SNFESTULOAD

54. For ILR records only. This field determines the student load that falls within the student numbers population for each further education record.

Value	Definition
SNFEPLANLEARNSTULOAD ÷ SNNUMFERECORDS	SNEXCL = 0 and SNLEVEL = FE and PLANLEARNHOURS ≠ BLANK
(SNHOURSPERAY ÷ 540) × 100	SNEXCL = 0 and SNLEVEL = FE and SNQUALHOURS ≠ BLANK and not above
10	SNEXCL = 0 and SNLEVEL = FE and not above
0	SNEXCL > 0 and SNLEVEL = FE
BLANK	Otherwise

SNTOTFESTULOAD

55. For ILR records only. This field calculates the sum of SNFESTULOAD across all records for the student. This is capped at a maximum of 150 per student.

SNTOTFEFTE

56. This field determines the total further education FTE.

Value	Definition
SNTOTFESTULOAD/100	SNSOURCE = ILR
0	SNSOURCE = DDB

Additional fields required to calculate total undergraduate higher education FTE relevant to TEF

SNREL_TEF

57. Identifies undergraduate provision relevant to TEF.

SNSOURCE = DDB

Value	Description	Definition
1	Provision in-scope for TEF	SNEXCL = 0 and
		(Z_LEVEL = C0000, C0001, C0002, C0003, C0004, C0005, C0006, C0007, C0008, C0009, H0003, H0004, H0005, H0006, H0007, H0008, H0012, H0015, H0018, H0019, H0020, I0001, I0004, I0008, I0009, I0012, I0013, J0000, J0001, J0002, J0003, J0004, J0005, J0006, J0007, J0010, J0011, J0012, M0002 or
		(Z_LEVEL = H0016, I0010 and PREREQUISITE ≠ 02))
0	Provision not in-scope for TEF	Otherwise

SNSOURCE = ILR

Value	Description	Definition
1	Provision in-scope for TEF	SNEXCL = 0 and OFSQAIM = FIRST, ENHANCED, FDBC, OTHL6_Q, OTHL6_CC, OTHL6_U, HND, DET, FOUDEG, DIPHE, DTLLS, OTHL5_Q, OTHL5_CC, OTHL5_U, HNC, CERTED, UNICERT, HIGHCERT, CTLLS, PTLLS, CET, OTHL4_Q, OTHL4_CC, OTHL4_U, OTHHE_Q, OTHHE_CC, OTHHE_U
0	Provision not in-scope for TEF	Otherwise

SNHEENGLEVSTULOAD_TEF

58. The total student load relevant to TEF for each engagement[6].

Value	Definition
SNHESTULOAD	SNSOURCE=DDB and SNREL_TEF=1
0	SNSOURCE=DDB and SNREL_TEF≠1
Sum of SNHESTULOAD across all provision relevant to TEF records (SNREL_TEF=1) for each unique combination of student (LEARNREFNUMBER) and level (SNLEVELNUM)	SNSOURCE=ILR

SNHEENGLEVSTULOADADJ_TEF

59. The capped student load relevant to TEF for each engagement[7].

$$SNHEENGLEVSTULOADADJ\_TEF = Min(SNHEENGLEVSTULOAD\_TEF, 150).$$

---

[6] The ILR data does not include the concept of an 'engagement'. For the purpose of calculating SNHEENGLEVSTULOAD we treat a level of study in the ILR data as analogous to an engagement. This means that for each row of data in the individualised file relating to ILR data, SNHEENGLEVLSTULOAD_TEF shows the student load relevant to TEF for that student at that level of study, regardless of whether the given row of data contains student load relevant to TEF

[7] The ILR data does not include the concept of an 'engagement'. For the purpose of calculating SNHEENGLEVSTULOADADJ_TEF we treat a level of study in the ILR data as analogous to an engagement.

SNTOTHESTULOAD_TEF

60. The total student load relevant to TEF per student, including any capping.

Value	Definition
Sum of SNHEENGLEVSTULOADADJ_TEF per student.	SNSOURCE=DDB
Sum of SNHEENGLEVSTULOADADJ_TEF across each unique combination of student and level.	SNSOURCE=ILR

SNTOTHEFTE_TEF

61. The total FTE relevant to TEF per student including any capping, calculated as SNTOTHESTULOAD_TEF÷100.

Rebuild instructions

62. The individualised file, SN23_PCO_XXXXXXXX_IND.csv, can be used to interrogate the data and our calculation of derived fields by filtering to certain groups of records. Full details of how to access and use an individualised file are given on the OfS website[8]https://www.officeforstudents.org.uk/data-and-analysis/supplying-data/working-with-individualised-files/.

63. Specific instructions to rebuild the student number summary workbook, SN23_PCO_XXXXXXXX.xlsx, are provided below.

64. A number of additional fields not described in the algorithms above have been added to the individualised file. These are listed in Table 5.

Table 5: Additional fields in the individualised file

Field name	Description
UKPRNY1Y2	The UKPRN of the provider that submitted the record. This might differ from UKPRN when a merger has taken place.
STUDENTIDTYPE	The type of student identifier field for the record. This will be SID for records from the Student record and LEARNREFNUMBER for records from the ILR.
STUDENTID	The SID or LEARNREFNUMBER for the record.
REBUILDLEVFLAG	A column provided to enable the figures in Table 1 and 2 to be rebuilt.
REBUILDHEFEFLAG	A column provided to enable the figures in Table 3 to be rebuilt.

65. REBUILDLEVFLAG and REBUILDHEFEFLAG are provided to allow you to easily de-duplicate the data in the individualised file. REBUILDHEFEFLAG is provided to select one row of data per student. REBUILDLEVFLAG is provided to select one row of data per student per level of higher education study.

66. The rebuild flags should not be used to determine which individual rows of data have been included or excluded from the calculation of student numbers. Instead, the SNEXCL field should be used to determine this.

Rebuilding the data in the summary file

67. The following instructions assume the user has the individualised file open in Microsoft Excel.

68. To apply filters to the column headings, highlight the headings. From the 'Data' tab in the toolbar select 'Filter'.

---

[8] See www.officeforstudents.org.uk/data-and-analysis/supplying-data/working-with-individualised-files/.

69. You will need to use the SUBTOTAL function with option 9 to sum the values in the relevant columns once the filters have been applied. Using the SUBTOTAL function instead of the SUM function will ensure that you only include those values that have been filtered to.

Table 1 Split of higher education provision by level of study (SNLEVELNUM)

70. To rebuild the figures in Table 1 of the workbook, filter to REBUILDLEVFLAG = 1 if you are using ILR data, otherwise skip this step. Next, sum SNHEENGLEVSTULOADADJ and divide by 100 to produce an FTE. This will give the total FTE in Table 1. To rebuild the figures for each level of study, filter SNLEVELNUM as required.

Table 2 Higher education provision relating to undergraduate levels of study relevant to TEF (SNREL_TEF)

71. To rebuild the figures in Table 2 of the workbook, first clear any filters. Filter to REBUILDLEVFLAG = 1. The subtotal for SNHEENGLEVSTULOADADJ_TEF, once divided by 100 to produce an FTE, will give the total FTE in Table 2.

Table 3 Split of further and higher education provision (SNLEVEL)

72. For ILR records only. The fields SNTOTFEFTE and SNTOTHEFTE are calculated per student. Where a student has more than one learning aim, SNTOTFEFTE and SNTOTHEFTE will be the sum of the further education and higher education FTE respectively across all rows of data.

73. To rebuild the figures in Table 3 of the workbook, filter to REBUILDHEFEFLAG = 1. The subtotals for SNTOTFEFTE and SNTOTHEFTE will give the totals for the further education and higher education FTE respectively. You do not need to filter by SNLEVEL as this has already been accounted for in generating REBUILDHEFEFLAG.